Aula 3: matrizes e camadas

13/11/2026, on-line. O que não couber entra na aula seguinte.

1. Revisão das aulas 1 e 2

de memória: o neurônio, o XOR, a perda, o passo do gradiente
$y = \text{degrau}(\vec{w} \cdot \vec{x} + b)$; $L = (1/N)\ \Sigma\ (y_i - t_i)^2$; $w \leftarrow w - \eta\ \partial L/\partial w$
o fio do curso, da aula 1 à 10; hoje, a 3

2. Uma camada, uma matriz

$\vec{h} = \text{degrau}(W\vec{x} + \vec{b})$: uma linha de W por neurônio; w_{kj}, linha k, coluna j
vídeo: a matriz vezes o vetor, $1 \cdot 2 + (-2) \cdot 1 + 0{,}5 = 0{,}5$; 2 por 2 vezes 2 dá 2
o XOR à mão: h_1 é o OU, h_2 é o E, $y = \text{degrau}(h_1 - h_2 - 0{,}5)$
as quatro entradas numa conta só: $Z = X\ W^T + \vec{b}$, X @ W.T + b
vídeo: a camada oculta leva os pontos a um plano em que uma reta separa
camada, camada oculta, MLP 2-2-1; hidden em Rumelhart, Hinton e Williams
quantos parâmetros: $n \cdot m + m$; 9, 12.730, 203.530, 68.353

3. A ativação

vídeo: a grade do plano sob W_1, W_2 e W_2W_1; a sigmoide entorta as retas
sem ativação, duas camadas valem uma: W_2W_1; conferido, $7 \cdot 10^{-15}$
com números: $W_2W_1 = (1\ 1)$, viés 5; (2; 1) dá 8 pelos dois caminhos
do degrau à sigmoide: $\sigma(z) = 1/(1 + e^{-z})$; $\sigma' = \sigma(1 - \sigma)$
vídeo: $\sigma(k\ z)$ vira degrau; a inclinação, 0,25 no zero; o degrau não treina
o XOR com a sigmoide: pesos vezes 1 erram, vezes 5 acertam, vezes 25, o degrau
a pergunta da aula 5: como achar os nove números sem escolher à mão?

4. O código

xor_a_mao.py na tela: a tabela das quatro entradas; duas lineares contra uma

5. Fechamento

3Blue1Brown, cálculo 4; Khan Academy, o Jacobiano
entrega: a passada para a frente para qualquer número de camadas
a seguir: aula 4, o Jacobiano; aula 5, a retropropagação